

Parameters of Two-Prover-One-Round Game and The Hardness of Connectivity Problems

Bundit Laekhanukit*

December 13, 2012

Abstract

Optimizing parameters of Two-Prover-One-Round Game (2P1R) is an important task in PCPs literature as it would imply a smaller PCP with the same or stronger soundness. While this is a basic question in PCPs community, the connection between the parameters of PCPs and hardness of approximations is sometime obscure to approximation algorithm community. In this paper, we investigate the connection between the parameters of 2P1R and the hardness of approximating the class of so-called connectivity problems, which includes as subclasses the survivable network design and (multi)cut problems. Based on recent development on 2P1R by Chan (ECCC 2011) and several techniques in PCPs literature, we improve hardness results of some connectivity problems that are in the form k^σ , for some (very) small constant $\sigma > 0$, to hardness results of the form k^c for some explicit constant c , where k is a connectivity parameter. In addition, we show how to convert these hardness into hardness results of the form \mathcal{D}^c , where \mathcal{D} is the number of demand pairs (or the number of terminals). Our results are as follows.

1. For the rooted k -connectivity problem, we have hardness of

$$\begin{cases} k^{1/2-\epsilon} & \text{on directed graphs.} \\ k^{1/10-\epsilon} & \text{on undirected graphs.} \\ \mathcal{D}^{1/4-\epsilon} & \text{on both directed and undirected graphs.} \end{cases}$$

This improves upon the best known hardness of k^σ by Cheriyan et al. (SODA 2012).

2. For the vertex-connectivity survivable network design problem, we have hardness of

$$\begin{cases} k^{1/6-\epsilon} & \text{on undirected graphs} \\ \mathcal{D}^{1/4-\epsilon} & \text{on both directed and undirected graphs.} \end{cases}$$

This improves upon the best known hardness of $\Omega(k^\sigma)$ by Chakraborty et al. (STOC 2008).

3. For the vertex-connectivity k -route cut problem on undirected graphs, we have hardness of

$$\begin{cases} k^{1/6-\epsilon} \\ \mathcal{D}^{1/4-\epsilon} \end{cases}$$

This improves upon the best known hardness of k^σ by Chuzhoy et al. (SODA 2012).

*School of Computer Science, McGill University, Montreal, QC, Canada. E-mail: blaekh@cs.mcgill.ca This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) grant no. 288334 and 429598, by European Research Council (ERC) Starting Grant no. 279352 and by Harold H Helm fellowship.

1 Introduction

Optimizing parameters of Two-Prover-One-Round Game (2P1R) is an important task in PCPs literature as it would imply a smaller PCP with the same or stronger soundness, which will in turn tighten hardness results for many optimization problems. While this is a basic question in PCPs community, the connection between the parameters of PCPs and hardness of approximations is sometime obscure to approximation algorithm community. In this paper, we investigate the connection between the parameters of 2P1R and the hardness of approximating the class of so-called connectivity problems, which includes as subclasses the survivable network design and (multi)cut problems.

Similar to 2P1R, a connectivity problem comes with several parameters, e.g., the number of vertices n , a connectivity parameter k , and the number of demand pairs \mathcal{D} . As these parameters are independent of each other, approximation algorithms for connectivity problems are usually designed by exploiting properties of the parameters, which means that the approximation ratios of the algorithms depend on these terms. By way of illustration, let consider a concrete example of the *rooted k -connectivity* problem on undirected graphs. In this problem, we are given an undirected graph $G = (V, E)$, a root vertex r and a set of terminals T ; the goal is to find a minimum-cost subgraph that has k openly (vertex) disjoint paths from the root vertex r to each terminal $t \in T$. For arbitrary k , the best known approximation ratio of this problem is $O(k \log k)$ by Nutov [24], and it was shown by Cheriyan, Laekhanukit, Naves and Vetta [7] that the dependence on k cannot be taken out because the problem does not admit $o(k^\sigma)$ -approximation, for some (very) small constant $\sigma > 0$, unless $P = NP$. However, when k is larger than the number of demands (or terminals) \mathcal{D} , a trivial \mathcal{D} -approximation algorithm does exist and yields a better approximation ratio than the $O(k \log k)$ -approximation algorithm. Moreover, the hardness result of Cheriyan et al. only holds when k is much smaller than \mathcal{D} . Thus, the approximability of the rooted k -connectivity problem on undirected graphs depends on two parameters: the connectivity k and the number of demands (terminals) \mathcal{D} , e.g., k is a constant independent of \mathcal{D} . Thus, to prove tighter approximation hardness of connectivity problems, we have to consider all the parameters involved.

Here two parameters of connectivity problems that we are interested in are the connectivity parameter k , which is the main focus in this paper, and the number of demand pairs \mathcal{D} . We consider 2P1R in its combinatorial form – the *label-cover* problem. In this problem, we are given a bipartite directed graph $G = (U, W, E)$, the set of labels (a.k.a., *alphabets*) L , and constraints which are functions on edges $\{\pi_e : e \in E\}$; the goal is to find an assignment of labels to each vertex that satisfies all the constraints. It is known that the hardness of the label cover problem depends on two parameters the *maximum degree* $\Delta(G)$ of G and the *alphabet-size* $|L|$. Thus, our goal is to investigate relationships between k , \mathcal{D} and $\Delta(G)$, $|L|$.

First, we consider the connectivity parameter k . The problems whose hardness depending on the parameter k are the rooted k -connectivity problem in both directed and undirected graphs, the *vertex-connectivity survivable network design* problem and the *vertex-connectivity k -route cut* problem. These problems have hardness of the form k^σ , where σ is a small constant that has not been calculated. (See [24, 7, 3, 11]). The common source of hardness of these problems is the label cover problem (a.k.a., 2P1R) with parallel repetition. The σ here involved with the constant loss in the exponent when boosting the hardness gap using parallel repetition. Estimating the value σ is not an easy task, and even if we can calculate this value, the constant is very small [17]. By studying the constructions of all these problems, we observe that the connectivity parameter k depends on the maximum degree and alphabet-size of a label cover instance. Thus, the simplest way in proving this hardness is to find an instance of the label cover problem whose maximum degree and alphabet-size are small comparing to the inverse of its soundness. Based on the recent developments in 2P1R [4, 22, 13, 2], we construct a label cover instance that satisfies the desired properties.

To be precise, we take a label cover instance of Chan [4] that has alphabet-size close to the inverse of its soundness. (Also, see the prior result by Khot and Safra [17].) Then we apply several reductions – the *right degree* reduction by Moshkovitz and Raz [22] and the *random sparsification* technique by Austrin, Khot and Safra [2]¹ to reduce the maximum degree of the instance. Hence, we have an instance with small degree and small alphabet-size, and we thus obtain the explicit exponent in the hardness of all the problems mentioned above.

Second, we consider the parameter \mathcal{D} , the number of demand pairs. The problems that we are interested in are the rooted k -connectivity problem on both directed and undirected graphs, the vertex-connectivity survivable network design problem and the vertex-connectivity k -route cut problem. By diving into the construction of these problems, we observe that some of the demand pairs are independent and thus can be merged. So, we have to partition the constraints (edges) of a label-cover instance so that they have no conflict after reducing to a connectivity problem. We observe that such partitioning can be done using *strong edge coloring*. To be precise, the strong edge coloring is a coloring of edges of G such that, for any two edges e, f with the same color, e and f share no endpoint, and G has no edge joining an endpoint of e to an endpoint of f . For example, edges $\{a, b\}$ and $\{c, d\}$ can have the same color if a, b, c, d are all distinct vertices, and G has none of the edges $\{a, c\}, \{a, d\}, \{b, c\}$ and $\{b, d\}$. It is known that a graph with maximum degree Δ has a strong edge coloring with $O(\Delta^2)$ colors. Thus, we can reduce the number of demands to be close to $\Delta(G)$, which is thus close to the inverse of its soundness.

Lastly, we would like to remark that we consider our results to be a survey paper that connects the parameters of 2PIR to the hardness of connectivity problems. All the techniques used in this paper are not new and have been used many times in literature. The right degree reduction was introduced by Moshkovitz and Raz in [22] and has been used in [14]. The random sampling technique was used in PCPs literature by Goldreich and Sudan in [15] and was recently used by Dinitz, Kortsarz and Raz in [13] to prove the hardness of the *basic k -spanner* problem. Also, it has been used to reduced the degree of an instance of the independent set problem by Austrin, Khot and Safra in [2]. Indeed, our work is inspired by the result of Moshkovitz and Raz [22] and the result of Dinitz et al. [13]. The graph coloring technique has been used to obtain approximation algorithms for the rooted k -connectivity problem in undirected graphs [9, 6, 24]. Here we show that such technique can be used to show the converse, i.e., the hardness of approximation. (Indeed, to best of our knowledge, the strong edge coloring has not been used in the previous literature.)

The connectivity problems considered in this paper are as follows.

The Rooted k -Connectivity Problem. In the *rooted k -connectivity* problem, we are given a directed or undirected graph $G = (V, E)$ on n vertices with cost c_e on each edge $e \in E$, a root vertex r , a set of terminals $T \subseteq V - \{r\}$ and a connectivity requirement k . The goal is to find a minimum-cost subgraph $G' = (V, E')$ of G such that G' has k *openly (vertex) disjoint* paths from r to each terminal $t \in T$. This problem has been studied intensively in [3, 9, 5, 10, 24, 25, 6, 7]. The rooted k -connectivity problem is a fundamental network design problem with vertex-connectivity requirements, and it lies at the bottom of the complexity hierarchy of the vertex-connectivity problems. In particular, the undirected rooted k -connectivity problem was shown to be a special case of the *subset k -connectivity* problem [20] and is clearly a special case of the *vertex-connectivity survivable network design* problem. It can be seen that the same relationships also apply for the case of directed graphs.

The rooted k -connectivity problem on both directed and undirected graphs admits a trivial $|T|$ -approximation algorithm, which can be done by applying a minimum-cost k -flow algorithm $|T|$ times, one for each ter-

¹ Indeed, we are first inspired by the result of Moshkovitz and Raz [22] and the result of Dinitz, Kortsarz and Raz [13]. However, due to a technical issue, we require a technique in [2], which was suggested by Siu On Chan.

minal. Non-trivial approximation algorithms for the rooted k -connectivity problem are known only for the undirected case, and the best known approximation ratio is $O(k \log k)$ by Nutov [24]; however, the approximation ratio surpasses that of the trivial algorithm only when $k > |T|$. On the negative side, Cheriyan, Laekhanukit, Naves and Vetta [7] recently showed that the rooted k -connectivity problem on both directed and undirected graphs are hard to approximate to within a factor of k^σ for some fixed $\sigma > 0$ (the constants σ are different in directed and undirected cases). However, the constants σ obtained are small and have not been explicitly calculated.

For the case of directed graphs, we give improved hardness of $k^{1/2-\epsilon}$ and $\mathcal{D}^{1/4-\epsilon}$ for the rooted k -connectivity problem, for any constant $\epsilon > 0$. (In fact, the $k^{1/2-\epsilon}$ -hardness of this problem can be derived from combining the result in [7] and [13].) For the case of undirected graphs, the hardness are $k^{1/10-\epsilon}$ and $\mathcal{D}^{1/4-\epsilon}$, for any constant $\epsilon > 0$, and this also gives the same bound for the hardness of the subset k -connectivity problem. (Note that the number of demand pairs is $\mathcal{D} = |T|$ for the rooted k -connectivity problem and $\mathcal{D} = |T|^2$ for the subset k -connectivity problem.)

The Vertex-Connectivity Survivable Network Design Problem. The *vertex-connectivity survivable network design* (VC-SNDP) problem is a generalization of the rooted k -connectivity problem. In this problem, we are given a directed or undirected graph $G = (V, E)$ on n vertices with a cost c_e on each edge e and a connectivity requirement $\text{req}(s, t)$ for each pair of vertices $s, t \in V$. A vertex s is called a *terminal* if there is a vertex t such that $\text{req}(s, t) > 0$, i.e., s is a terminal if it has a positive connectivity requirement; the set of terminals is denoted by T . The only known non-trivial approximation algorithm for this problem due to the work of Chuzhoy and Khanna [10] has an approximation ratio of $O(k^3 \log |T|)$, and the best known hardness is k^σ , for some (very) small constant $\sigma > 0$, due to Chakrabarty, Chuzhoy and Khanna [3]. We give an improved hardness of $k^{1/6-\epsilon}$ and $\mathcal{D}^{1/4-\epsilon}$ for VC-SNDP, for any constant $\epsilon > 0$.

The Vertex-Connectivity k -Route Cut Problem. In the *vertex-connectivity k -route cut* (VC- k -RC) problem, we are given an undirected graph $G = (V, E)$ on n vertices with a cost c_e on each edge $e \in E$, a set of source-sink pairs $\{(s_1, t_1), (s_2, t_2), \dots, (s_{\mathcal{D}}, t_{\mathcal{D}})\} \subseteq V \times V$ and a connectivity parameter k . The goal is to find a minimum-cost subset $E' \subseteq E$ of edges such that $G - E'$ has no k openly disjoint s_i, t_i -paths for every source-sink pairs s_i, t_i . The best known approximation guarantee for this problem is $O(\mathcal{D} \cdot k)$ due to the work of Chuzhoy, Makarychev, Vijayaraghavan and Zhou [11], and the best known hardness is $O(k^\sigma)$, for some (very) small constant $\sigma > 0$. The approximation ratio is slightly better when we turn to a bi-criteria approximation algorithm. Chuzhoy et al. showed that there is an algorithm that guarantees to find a solution $E' \subseteq E$ with cost at most $O(\lambda k \log^{2.5} \mathcal{D} \log \log \mathcal{D})$ times the optimal, where λ is the maximum number of demand pairs in which any terminal participates, and E' cuts at least $k/2$ -routes, i.e., $G - E'$ has no $k/2$ openly s_i, t_i -paths for all i . In this paper, we show that at least one of the two terms k and \mathcal{D} cannot be taken out. Precisely, we show that it is hard to approximate VC- k -RC to within a factor of $k^{1/6-\epsilon}$ and $\mathcal{D}^{1/4-\epsilon}$, for any constant $\epsilon > 0$.

Our hardness results are summarized in Table 1.

2 Preliminaries

We use standard graph terminologies as in [12]. Let $G = (V, E)$ be any graph. For any vertex $v \in V$, the degree of v in G is denoted by $\deg_G(v)$. The maximum (resp., minimum) degree of G , denoted by $\Delta(G)$ (resp., $\delta(G)$), is the maximum (resp., minimum) degree over all vertices of G . If we consider more than one graph, then we denote the set of vertices and edges of G by $V(G)$ and $E(G)$, respectively.

| Problem | Graphs | In terms of k ($k < \mathcal{D}$) | In terms of \mathcal{D} ($k \geq \mathcal{D}$) |
|--------------------------|------------|---------------------------------------|--|
| Rooted k -Connectivity | Directed | $k^{1/2-\epsilon}$ | $\mathcal{D}^{1/4-\epsilon}$ |
| | Undirected | $k^{1/10-\epsilon}$ | $\mathcal{D}^{1/4-\epsilon}$ |
| Subset k -Connectivity | Undirected | $k^{1/10-\epsilon}$ | $\mathcal{D}^{1/4-\epsilon}$ |
| VC-SNDP | Undirected | $k^{1/6-\epsilon}$ | $\mathcal{D}^{1/4-\epsilon}$ |
| VC- k -Route Cut | Undirected | $k^{1/6-\epsilon}$ | $\mathcal{D}^{1/4-\epsilon}$ |

Table 1: The table summarizes our hardness results, which hold for any $\epsilon > 0$.

By a bipartite directed graph, we mean a directed graph $G = (U, W, E)$ such that every arc is directed from U to W , i.e., an arc of G is of the form (u, w) , where $u \in U$ and $w \in W$. We call vertices in U *left vertices* and vertices in W *right vertices*. Since each left (resp., right) vertex of G has no incoming (resp., outgoing) arc, we abuse the term “degree” to mean indegree (resp., outdegree) of left (resp., right) vertices of G . By the *maximum (resp., minimum) left degree* of G , denoted by $\Delta_{\text{left}}(G)$ (resp., $\delta_{\text{left}}(G)$), we mean the maximum (resp., minimum) degree of left vertices of G . Similar notations are used for right vertices. Thus, $\Delta_{\text{right}}(G)$ (resp., $\delta_{\text{right}}(G)$) denotes the maximum (resp., minimum) degree of right vertices of G . We use a similar notations for the average degree of G . The average degree of G is denoted by $\Delta^{\text{avg}}(G)$, and the average left and right degree of G are denoted by $\Delta_{\text{left}}^{\text{avg}}(G)$ and $\Delta_{\text{right}}^{\text{avg}}(G)$, respectively. We say that G is *left (resp., right) regular* if every left (resp., right) vertex of G has the same degree. If G is both left and right regular with degree d_1 and d_2 , then we say that G is (d_1, d_2) -*regular*. If G is clear in the context, then we will omit G , e.g., we may write $\deg_G(v)$ as $\deg(v)$ and write $\Delta(G)$ as Δ .

By a *matching* M of a (directed) graph G , we mean a set of edges (resp., arcs) such that no two edges (resp., arcs) in M share an endpoint, and by *induced matching* I in G , we mean a matching such that no edge (resp., arc) in G joins endpoints of edges (resp., arcs) in I . Thus, a subgraph of G induced by such I is also a matching. A *strong edge coloring* of G is a partition E_1, E_2, \dots, E_ℓ of sets of edges (resp., arcs) of G such that each E_i is an induced matching in G . The smallest number ℓ such that G has an ℓ -strong edge coloring is called the *strong chromatic index* of G , denoted by $\chi'_S(G)$.

All of our hardness results come from the same source, *the label cover problem* (a.k.a, 2P1R). Hence, we devote the next section to discuss the label cover problem.

2.1 The Label Cover Problem

The *(maximum) label cover problem* (*the projection game*) is defined as follows. We are given a directed bipartite graph $G = (U, W, E)$ on n vertices, two sets of labels (a.k.a, alphabets) L_1 (for vertices in U) and L_2 (for vertices in W), and a *constraint* π_e on each arc e , which is a *projection*² $\pi_e : L_1 \rightarrow L_2$. A *labeling* (f_1, f_2) is a pair of functions $f_1 : U \rightarrow L_1$ and $f_2 : W \rightarrow L_2$ assigning a label to each vertex of U and W , respectively. We say that (f_1, f_2) *covers* an arc $(u, w) \in E$ if $\pi_e(f_1(u)) = f_2(w)$. The goal in the maximum label cover problem is to find a labeling that maximizes the number of arcs covered. For notational convenience, we shall denote an instance of the label cover problem by $(G = (U, W, E), \{\pi_e : e \in E\}, L_1, L_2)$.

The gap version of the maximum label cover problem is the problem of deciding whether a given instance of the maximum label cover problem is one of the following two cases:

² The constraints of the label cover problem can be relations instead of projections; however, here we define the label cover problem as the projection game.

- YES-INSTANCE: There is an labeling covering at least $(1 - \epsilon)$ fraction of all the arcs.
- NO-INSTANCE: There is no labeling covering more than γ fraction of all the arcs.

We call $1 - \epsilon$ and γ the *completeness* and the *soundness* of the label cover instance, respectively. If $\epsilon = 0$, then we say that a (gap) label cover instance has *perfect completeness*; otherwise, we say that an instance has *imperfect completeness*. It can be seen that NP-hardness of the gap version of the label cover problem implies the hardness of the maximum one. Thus, we shall abuse the term “maximum label cover” to also mean the gap label cover problem.

For our purpose, we need a minimization version of the label cover problem, which can be defined by allowing each vertex to have more than one label, and the goal is to minimize the total cost of labels used over all vertices. To be precise, we define the *minimum-cost label cover* problem to be the weighted counter part of the maximum label cover problem. The minimum label cover problem was defined in [1], and it has an equivalent form known as the *Min-Rep* problem as defined in [18]. The input of this problem is the same as that of the maximum label cover problem except that we also have a cost c_1 on each label $a \in L_1$ and a cost c_2 on each label $b \in L_2$. The labeling is relaxed as a pair of functions (f_1, f_2) , where $f_1 : U \rightarrow 2^{L_1}$ and $f_2 : W \rightarrow 2^{L_2}$, i.e., we are allowed to assign more than one labels to each vertex. A labeling (f_1, f_2) covers an arc $e = (u, w)$ if there are labels $a \in f_1(u)$ and $b \in f_2(w)$ such that $\pi_e(a) = b$. The goal in the minimum-cost label cover problem is to find a labeling (f_1, f_2) that covers all the arcs and minimizes the cost $c(f_1, f_2) = \sum_{u \in U} c_1 \cdot |f_1(u)| + \sum_{w \in W} c_2 \cdot |f_2(w)|$.

Note that there is a standard technique that transforms the hardness of the maximization version of the label cover problem to the minimum-cost version. (See Appendix A for more detail.) Thus, it suffices to consider the maximum label cover problem.

3 Relationships between Label Cover and with Connectivity Problems

Here we show the relationships between the parameters of the label cover problem (2P1R) with the hardness of approximating connectivity problems.

First, we survey relationships between the hardness in terms of the connectivity parameter k of the connectivity problems and the parameters of the minimum-cost label cover problems.

Theorem 1 ([7, 3, 11]). *Given an instance $(G = (U, W, E), \{\pi_e : e \in E\}, L_1, L_2, c_1, c_2)$ of the minimum-cost label cover problem, there are polynomial-time approximation preserving reductions that output*

- *An instance of the rooted k -connectivity problem on directed graphs with $k = \Delta(G)$.*
- *An instance of the rooted k -connectivity problem on undirected graphs with $k = O(\Delta(G)^3 \cdot \max\{|L_1|, |L_2|\} + \Delta(G)^4)$.*
- *An instance of the vertex-connectivity survivable network design problem on undirected graphs with a maximum requirement $k = O(\Delta(G) \cdot \max\{|L_1|, |L_2|\} + \Delta(G)^2)$.*
- *An instance of the vertex-connectivity k -route cut problem on undirected graphs with $k = O(\Delta(G) \cdot \max\{|L_1|, |L_2|\})$.*

The hardness in terms of the connectivity parameter k can be transformed into hardness in terms of the number of demand pairs \mathcal{D} . The parameter that involves with hardness in this term is the degree of the label cover instance. We claim that, for each of the problems we consider, two demand pairs (s_1, t_1) and (s_2, t_2) are *independent* if and only if they come from two different constraints (arcs) (u_1, w_1) and (u_2, w_2) of the label cover instance such that (u_1, w_1) and (u_2, w_2) forms an induced matching, which thus can have

the same “strong edge color”. So, we can partition the arcs of the label cover instance using strong edge coloring and merge source-sink pairs with the same color. Thus, we have the following theorem.

Theorem 2. *For each of the following problems, say Π ,*

- *The rooted k -connectivity problem on directed graphs,*
- *The rooted k -connectivity problem on undirected graphs,*
- *The vertex-connectivity survivable network design problem on undirected graphs,*
- *The vertex-connectivity k -route cut problem on undirected graphs,*

there is a polynomial-time reduction that, given an instance $(G = (U, W, E), \{\pi_e : e \in E\}, L_1, L_2, c_1, c_2)$ of the minimum-cost label cover problem, outputs an instance of the problem Π with the number of demand pairs $\mathcal{D} = 2\Delta(G)^2$.

See Appendix C, D, E, and F for the full proofs and discussions.

As we will show in the next section, the label cover instance of Chan [4] can be modified so that it has degree close to the inverse of its soundness. (See Theorem 6). We apply a standard technique to transform the hardness of the maximization version of the label cover problem to the minimum-cost version. (See Appendix A for more detail.) Then we have the following theorem.

Theorem 3. *For any constants $q > 0$ and $\epsilon > 0$, given an instance $(G = (U, W, E), \{\pi_e : e \in E\}, L_1, L_2)$ of the maximum label cover such that $|L_1|, |L_2| \leq q^2$, $\Delta(G) = \Theta(q)$ and $\Delta^{\text{avg}}(G) = \Theta(q)$, unless $\text{NP} = \text{ZPP}$, it is hard to approximate the minimum-cost label cover problem to within a factor of $q^{1/2-\epsilon}$.*

By substituting the bound in Theorem 3 to Theorem 1 and Theorem 2, we prove the results in Table 1.

4 Modifying The Label Cover Instance

In this section, we show how to construct a label cover instance with strong soundness, small degree and small alphabet-size. In particular, we prove the following lemma.

Lemma 4. *Let $q > 0$ be a constant. There is a randomized polynomial-time algorithm that reads as input an instance $(G = (U, W, E), \{\pi_e : e \in E\}, L_1, L_2)$ of the maximum label-cover problem with the following properties:*

- *The alphabet-size is $\max\{|L_1|, |L_2|\} \leq q$.*
- *The graph G has regular left degree $\Delta_{\text{left}} = \text{poly}(q)$.*
- *The completeness is $1 - \epsilon$, for any $\epsilon > 0$.*
- *The soundness is $\gamma = 1/\text{poly}(q)$.*

outputs an instance $(G' = (U', W', E'), \{\pi_e : e \in E\}, L_1, L_2)$ of the maximum label-cover problem with completeness $1 - \epsilon$ and soundness $\gamma' = \Theta(\gamma)$ and $\Delta(G') \leq O((1/\gamma) \log(1/\gamma))$.

The following theorem is due to the work of Chan [4].

Theorem 5 ([4]). *For any constants $q > 0$ and $\epsilon > 0$, given an instance $(G = (U, W, E), \{\pi_e : e \in E\}, L_1, L_2)$ of the maximum label cover such that $|L_1|, |L_2| \leq q^2$ and $\Delta_{\text{left}}(G) = q$, it is NP-hard to distinguish between the following two cases.*

- **YES-INSTANCE:** *There is an labeling covering at least $(1 - \epsilon)$ fraction of all the arcs.*

- NO-INSTANCE: *There is no labeling covering more than $O(\log q/q)$ fraction of all the arcs.*

Thus, by invoking Lemma 4, we have the following theorem.

Theorem 6. *For any constants $q > 0$ and $\epsilon > 0$, given an instance $(G = (U, W, E), \{\pi_e : e \in E\}, L_1, L_2)$ of the maximum label cover such that $|L_1|, |L_2| \leq q^2$, $\Delta(G) = \Theta(q \log q)$ and $\Delta^{\text{avg}}(G) = \Theta(q \log q)$, unless $\text{NP} = \text{ZPP}$, it is hard to distinguish between the following two cases.*

- YES-INSTANCE: *There is an labeling covering at least $(1 - \epsilon)$ fraction of all the arcs.*
- NO-INSTANCE: *There is no labeling covering more than $O(\log q/q)$ fraction of all the arcs.*

So, we devote the remaining part of this section to prove Lemma 4. We have four steps. First, we take a basic instance, which is a label cover instance with strong soundness and have a regular left-degree as in Theorem 5. We apply the right degree reduction to make a (d_1, d_2) -regular instance. Then we make copies of left vertices so that both sides have the same number of vertices and thus have regular degree. Finally, we apply a random sparsification to reduce the maximum degree of a label cover graph to be $O(q \log q)$, where $\gamma = 1/\text{poly}(q)$ is the soundness of the label cover instance.

4.1 Basic Instance

We take an instance $(G = (U, W, E), \{\pi_e : e \in E\}, L_1, L_2)$ of the maximum label cover problem with properties as stated in Lemma 4. That is,

- The alphabet-size is $\max\{|L_1|, |L_2|\} \leq q$.
- The graph G has regular left degree $D = \text{poly}(q)$.
- The completeness is $1 - \epsilon$, for any $\epsilon > 0$.
- The soundness is $\gamma = 1/\text{poly}(q)$.

An instance of the maximum label cover problem that satisfies the above properties are that constructed by Chan in [4] and by Khot and Safra in [17]. Note that due to the size of the construction, the former result applies for any constant $q > 0$ while the latter result in [17] applies for all primes $5 \leq q \leq \text{polylog}(N)$, where N is the size of the label cover instance. More precisely, the result in [17] also applies for $q = \text{polylog}(N)$ under the hardness assumption $\text{NP} \subsetneq \text{DTIME}(2^{\text{polylog} n})$.

4.2 Making An Instance (d_1, d_2) -Regular

The basic instance discussed in the previous chapter is the bipartite graph $G = (U, V, E)$ that is left-regular but not right-regular. To make an instance of the maximum label cover instance regular, we apply the *right degree reduction* introduced by Moshkovitz and Raz [22]. (Also, see [14].) In short, the right degree reduction makes the right degree of a label cover instance regular while almost preserves the soundness. It is not hard to see that the reduction preserves the completeness as well. (See Appendix B for more detail.)

Lemma 7 (Right Degree Reduction [22]). *There exists a polynomial-time reduction that, given a parameter d and a maximum label cover instance $(G = (U, W, E), L_1, L_2, \{\pi_e : e \in E\})$ with completeness $1 - \epsilon$ and soundness γ , where G has regular left degree Δ_{left} , outputs a maximum label cover instance $(G' = (U', W', E'), L_1, L_2, \{\pi_e\}_{e \in E'})$ with regular left degree $d \cdot \Delta_{\text{left}}$, regular right degree d , completeness $1 - \epsilon$ and soundness $\gamma + O(1/\sqrt{d})$.*

We choose a parameter $d = 1/\gamma$ and apply the right degree reduction on G . Thus, we have an instance $(\widehat{G} = (U^k, \widehat{W}, \widehat{E}), L_1, L_2, \{\pi_e\}_{e \in \widehat{E}})$ of the maximum label cover problem in which \widehat{G} is (dD, d) -regular, where D is the left-degree of G , as desired.

4.3 Making (d_1, d_2) -Regular Instance Δ -Regular

Take an instance \widehat{G} of the maximum label cover problem as discussed in the previous section. Now, we want to make the (dD, d) -regular graph \widehat{G} a dD -regular graph. To do so, we replace each left vertex u of G by D vertices u_1, u_2, \dots, u_D and we add an arc (u_i, w) with a constraint $\pi_{u_i, w} = \pi_{u, w}$, for each arc (u, w) of \widehat{G} . This results in a graph G^{reg} which is dD -regular because the degree of each right vertex increases by a factor of D while the degree of each left vertex remains the same. Observe that the reduction preserves completeness because each edge of \widehat{G} has exactly D copies in G^{reg} . Now, consider the soundness. Take any labeling (f_1, f_2) of G^{reg} . We construct a labeling (\hat{f}_1, f_2) of \widehat{G} by assigning $\hat{f}_1(u) = f_1(u_i)$, where $f_1(u_i)$ is a labeling that covers the maximum number of arcs of G^{reg} incident to u_i given that f_2 is fixed. If (f_1, f_2) covers more than γ fraction of arcs of G^{reg} , then (\hat{f}_1, f_2) will cover more than γ fraction of arcs of \widehat{G} as well by the choice of $\hat{f}_1(u)$. Therefore, the reduction preserves both completeness and soundness, and the resulting bipartite graph is Δ -regular, where $\Delta = dD$.

4.4 Reducing Degree via Random Sparsification

Now, we take a dD -regular label cover instance from the previous section. We apply a random sparsification technique to reduce the “average degree” of the label cover instance to almost match the inverse of its soundness. Then we throw away vertices with large degree so that the graph has degree within the desired bound, the inverse of the soundness.

4.4.1 Sparsifying The Graph

First, we will sparsify the graph \widehat{G} . The reduction takes as input a regular-degree instance of the maximum label cover problem and outputs an instance whose bipartite graph has small average degree. To be precise, the input of our reduction is an instance $(G = (U, W, E), \{\pi_e : e \in E\}, L_1, L_2)$ of the maximum label cover problem with regular degree Δ , completeness $1 - \epsilon$ and soundness γ . Then it constructs a graph $G' = (U, W, E')$ from G by randomly and independently picking each arc of G with probability $\rho = \gamma^{-1} \log(\max\{|L_1|, |L_2|\}) / \Delta(G)$.

Intuitively, since we sample arcs of G with the same probability ρ , the resulting graph G' should have degree approximately $O(\rho\Delta)$, and for any labeling (f_1, f_2) , the fraction of arcs in G' that (f_1, f_2) covers is approximately the same as that it covers in G . The next theorem shows that the random sparsification (almost) preserves completeness and soundness of the original instance. Moreover, the average degree of the output instance is exactly $\Delta^{\text{avg}}(G) = \gamma^{-1} \log(\max\{|L_1|, |L_2|\})$.

Lemma 8. *Suppose the random sparsification algorithm takes as input an instance $(G = (U, W, E), \{\pi_e : e \in E\}, L_1, L_2)$ of the maximum label cover problem with regular degree $\Delta(G)$, completeness $1 - \epsilon$ and soundness γ , where $0 < \epsilon, \gamma < 1$. Then it outputs with high probability an instance $G' = (U, W, E', \{\pi_e : e \in E'\}, L_1, L_2)$ of the maximum label cover problem with completeness $1 - 4\epsilon$, soundness 8γ and the average degree of G' is $\Delta^{\text{avg}}(G') = \gamma^{-1} \log(\max\{|L_1|, |L_2|\})$.*

Proof. Throughout, let $n = |U| + |W|$ denote the number of vertices of G .

Completeness: Suppose there is a labeling (f_1, f_2) covering $(1 - \epsilon)$ fraction of arcs in G . We will show that (f_1, f_2) covers at least $1 - 2\epsilon$ fraction of arcs in G' .

Let $X = \sum_{e \in E} X_e$ be the number of arcs covered by the labeling (f_1, f_2) , where X_e is an indicator random variable such that $X_e = 1$ if an arc e is covered by the labeling (f_1, f_2) and $X_e = 0$ otherwise.

Then the expected number of arcs not covered by (f_1, f_2) is

$$\mathbf{E}[|E| - X] = \epsilon|E| \cdot \frac{\gamma^{-1}}{\Delta(G)} \log(\max\{|L_1|, |L_2|\}) = \frac{\epsilon n}{2} \cdot \gamma^{-1} \log(\max\{|L_1|, |L_2|\})$$

The last equation follows since $|E| = \Delta(G)|U| = \Delta(G)|W|$. By Chernoff's bound, we have

$$\Pr[|E| - X > \epsilon n \gamma^{-1} \log(\max\{|L_1|, |L_2|\})] < \exp\left(-\frac{\epsilon n}{6} \cdot \gamma^{-1} \log(\max\{|L_1|, |L_2|\})\right) \leq 2^{-\Omega(n)}$$

Now, consider the expected number of arcs in G' . We have

$$\mathbf{E}[|E'|] = |E| \cdot \frac{\gamma^{-1} \log(\max\{|L_1|, |L_2|\})}{\Delta(G)} = \frac{n}{2} \cdot \gamma^{-1} \log(\max\{|L_1|, |L_2|\}).$$

Thus, by Chernoff's bound,

$$\Pr[|E'| < \frac{n}{4} \gamma^{-1} \log(\max\{|L_1|, |L_2|\})] < \exp\left(-\frac{n}{16} \gamma^{-1} \log(\max\{|L_1|, |L_2|\})\right) \leq 2^{-\Omega(n)}$$

By union bound, with high probability, (f_1, f_2) covers at least $(1 - 4\epsilon)$ fraction of arcs in G' .

Soundness: Suppose there is no labeling (f_1, f_2) covering more than γ fraction of arcs in G . We will show that there is also no labeling (f_1, f_2) covering more than 4γ fraction of arcs in G' .

Fix any labeling (f_1, f_2) . Let $X = \sum_{e \in E} X_e$ be the number of arcs covered by the labeling (f_1, f_2) , where X_e is an indicator random variable such that $X_e = 1$ if an arc e is covered by the labeling (f_1, f_2) and $X_e = 0$ otherwise. Then the expected number of arcs satisfied by (f_1, f_2) is

$$\mathbf{E}[X] = \gamma|E| \cdot \frac{\gamma^{-1} \log(\max\{|L_1|, |L_2|\})}{\Delta(G)} = \frac{n}{2} \cdot \log(\max\{|L_1|, |L_2|\}).$$

Thus, by Chernoff's bound, we have

$$\Pr[X > 2n \log(\max\{|L_1|, |L_2|\})] < \exp\left(\frac{9}{6} n \log(\max\{|L_1|, |L_2|\})\right) \leq \left(\frac{1}{\max\{|L_1|, |L_2|\}}\right)^{3n/2}$$

Since there are $|L_1|^{n/2} |L_2|^{n/2} \leq (\max\{|L_1|, |L_2|\})^n$ possible labellings, by union bound, we have that with probability $2^{-\Omega(n)}$, the labeling (f_1, f_2) covers at most $2n \log(\max\{|L_1|, |L_2|\})$ arcs. By the proof for the case of **Yes-Instance**, we have that G' has at most $(n/4)\gamma^{-1} \log(\max\{|L_1|, |L_2|\})$ arcs with probability $2^{-\Omega(n)}$. Thus, the labeling (f_1, f_2) covers at most 8γ fraction of the arcs with high probability. This completes the proof. \square

So, we can sparsify the instance $(\widehat{G} = (U, W, \widehat{E}), \{\pi_e : e \in \widehat{E}\}, L_1, L_2)$ from Section 4.3 to obtain an instance $(\widehat{G}^{avg} = (U, W, \widehat{E}^{avg}), \{\pi_e : e \in \widehat{E}^{avg}\}, L_1, L_2)$ such that \widehat{G}^{avg} has average degree $\Delta^{avg}(G') = \gamma^{-1} \log(\max\{|L_1|, |L_2|\})$.

4.4.2 Removing Vertices with Large Degree

The graph \hat{G}^{avg} obtained from the previous step has average degree to be within the desired bound. However, some vertices may still have *large degree*, i.e., their degree are larger than $2\gamma^{-1} \log(\max\{|L_1|, |L_2|\})$. To make the graph to have degree within the desired bound, we remove all the large degree vertices from the graph \hat{G}^{avg} . This results in a graph $\hat{G}^{bound} = (U', V', E')$ with $\Delta(\hat{G}^{bound}) \leq 2\gamma^{-1} \log(\max\{|L_1|, |L_2|\})$.

By Chernoff's bound, the probability that a vertex $v \in U \cup W$ has large degree is

$$\Pr[\deg_{\hat{G}^{avg}}(v) > 2\gamma^{-1} \log(\max\{|L_1|, |L_2|\})] < \exp\left(-\frac{1}{3}\gamma^{-1} \log(\max\{|L_1|, |L_2|\})\right) \leq 2^{-(1/3)\gamma^{-1}}$$

Let X_v be an indicator variable such that $X_v = 1$ if v has large degree and $x_v = 0$ otherwise. Then we have $\mathbf{E}[\sum_{v \in V \cup W} X_v] = \sum_{v \in V \cup W} \mathbf{E}[X_v] \leq 2^{-(1/3)\gamma^{-1}} |V \cup W|$.

By Markov's inequality, we have $\Pr\left[\sum_{v \in V \cup W} X_v > 2^{-(1/3)\gamma^{-1}+1} |V \cup W|\right] < \frac{1}{2}$.

Thus, with probability $1/2$ we remove at most $2^{-(1/3)\gamma^{-1}} |V \cup W|$ vertices of \hat{G}^{avg} ; we call this a *probability of success*. We can repeat the process $O(\log n)$ times, where $n = |U| + |W|$, to increase the probability of success to $1 - 1/\Omega(n)$. This does not effect the success probability of the random sparsification step because the probability of success of the random sparsification step is very high, say $1 - 1/2^{\Omega(n)}$.

The Size of Construction: As above, with high probability, the graph \hat{G}^{bound} has at least $(1 - 1/2^{n/3})n$ vertices. For the number of arcs, we may assume that all vertices removed have degree $dD = \text{poly}(\gamma^{-1})$, where dD is the (regular) degree of the graph \hat{G} . Note that γ^{-1} is smaller than $O(\log n)$. So, with high probability, the number of arcs of \hat{G}^{bound} is at least $|E(\hat{G}^{avg})| - 2^{-n/3}dD(|U| + |W|) \geq (1 - 2^{-n/6})|E(\hat{G}^{avg})|$.

Completeness: Suppose there is a labeling (f_1, f_2) covering $(1 - \epsilon)$ fraction of the arcs of \hat{G}^{avg} . We will show that (f_1, f_2) covers at least $(1 - 2\epsilon)$ fraction of the arcs of \hat{G}^{bound} .

We may assume that arcs incident to vertices removed from \hat{G}^{avg} are covered by (f_1, f_2) , and each vertex removed has degree $dD = \text{poly}(\gamma^{-1})$, which is the maximum degree of \hat{G} . Thus, the number of arcs of \hat{G}^{bound} covered by (f_1, f_2) is at least $(1 - \epsilon)|E(\hat{G}^{avg})| - 2^{-n/3}dD(|U| + |W|) \geq (1 - 2\epsilon)|E(\hat{G}^{avg})|$.

The last inequality follows because ϵ is a constant. Therefore, in this case, there is a labeling that covers at least $(1 - 2\epsilon)$ fraction of the arcs of \hat{G}^{bound} .

Soundness: Suppose there is no labeling (f_1, f_2) that covers more than γ fraction of the arcs of \hat{G}^{avg} . We will show that there is no labeling (f_1, f_2) that covers more than 2γ fraction of the arcs of \hat{G}^{bound} .

Consider a labeling (f_1, f_2) of \hat{G}^{avg} that covers $\gamma' \leq \gamma$ fraction of arcs of \hat{G}^{bound} . We construct a labeling (f'_1, f'_2) of \hat{G}^{bound} by assigning $f'_1(u) = f_1(u)$ (resp., $f'_2(w) = f_2(w)$) for each vertex $u \in U'$ (resp., $w \in W'$) in \hat{G}^{bound} . We may assume the worst case that all vertices removed have degree dD and arcs incident to them are not covered by (f_1, f_2) . So, (f'_1, f'_2) still covers $\gamma'|E(\hat{G}^{avg})|$ arcs of \hat{G}^{bound} , but the number of arcs in \hat{G}^{bound} is smaller than that of \hat{G}^{avg} . By the analysis of the construction size, with high probability, \hat{G}^{bound} has at least $(1 - 2^{-n/6})|E(\hat{G}^{avg})|$ arcs. Thus, (f_1, f_2) covers at most $\frac{\gamma'}{1 - 2^{-n/6}} \leq 2\gamma' \leq 2\gamma$ arcs of \hat{G}^{bound} .

The first inequality follows because $1 - 2^{-n/6} \geq 1/2$ for large enough n . Therefore, there is no labeling covering more than 2γ fraction of arcs of \hat{G}^{bound} . Moreover, this happens with high probability.

This completes the proof of Lemma 4.

5 Getting Hardness in Terms of \mathcal{D}

In this section, we discuss how to obtain the hardness in terms of demand pairs. We will give an example of the hardness of the rooted k -connectivity problem on directed graphs. The next theorem is implicit in the construction of Cheriyan et al. [7].

Theorem 9 (Implicit in [7]). *There is a polynomial-time approximation preserving reduction such that, given an instance $(G = (U, W, E), \{\pi_e : e \in E\}, L_1, L_2, c_1, c_2)$ of the minimum-cost label cover problem, outputs an instance (\widehat{G}, c, r, T) of the rooted k -connectivity problem on directed graphs with $k = O(\delta(G))$, where \widehat{G} is a directed graph, c is a cost function, r is a root vertex and T is a set of terminals. Moreover, the reduction has the following properties:*

- Each terminal $t_{i,j} \in T$ corresponds to an arc $(u_i, w_j) \in E(G)$.
- The graph \widehat{G} can be partitioned into $\widehat{G} = \bigcup_{t_{i,j} \in T} \widehat{G}_{i,j}$, where $\widehat{G}_{i,j}$ is the union of all $r, t_{i,j}$ -paths in \widehat{G} .
- For any two partitions $\widehat{G}_{i,j}$ and $\widehat{G}_{i',j'}$ of \widehat{G} , where $i \neq i'$ and $j \neq j'$, there is a path from $\widehat{G}_{i,j}$ to $\widehat{G}_{i',j'}$ (resp., from $\widehat{G}_{i',j'}$ to $\widehat{G}_{i,j}$) if and only if the label cover graph G has an arc $(u_i, w_{j'})$ (resp., $(u_{i'}, w_j)$).

The full discussions are provided in Appendix C, and the discussions for other problems are discussed in Appendix D, E, and F.

Our goal is to reduce the number of terminals by merging some terminals of the instance (\widehat{G}, c, r, T) of the rooted k -connectivity problem on directed graphs as in Theorem 9. However, if we merge terminals $t_{i,j}$ and $t_{i',j'}$ such that $\widehat{G}_{i,j}$ and $\widehat{G}_{i',j'}$ share some non-root vertex, then this will cause us some problems. For example, we might have some “free path” formed by concatenating an $r, t_{i,j}$ -path and an $r, t_{i',j'}$ -path, or we might not have enough openly disjoint paths to satisfy the connectivity requirement. Thus, we have to ensure that no two terminals that we merge share a non-root vertex in the graphs $\widehat{G}_{i,j}$ ’s.

Observe that if the label cover graph G has no arc joining (u_i, w_j) and $(u_{i'}, w_{j'})$, where $i \neq i'$ and $j \neq j'$, then the graphs $\widehat{G}_{i,j}$ and $\widehat{G}_{i',j'}$ share no non-root vertex. In other words, if (u_i, w_j) and $(u_{i'}, w_{j'})$ form “an induced matching” in G , then we can merge terminals $t_{i,j}$ and $t_{i',j'}$. Hence, we can partition arcs of G into induced matching by applying “strong edge coloring”, and the number of partition of arcs we obtain is at most $2\Delta(G)^2$. Thus, we can merge terminals in T into $2\Delta(G)^2$ terminals. Applying Theorem 3, we have the hardness of $|T|^{1/4-\epsilon} = \mathcal{D}^{1/4-\epsilon}$, for any $\epsilon > 0$ as claimed.

Acknowledgment. We thank Adrian Vetta, Joseph Cheriyan, Guylain Naves, Parinya Chalermsook, Danupon Nanongkai and Siu On Chan for useful comments and discussions.

References

- [1] Sanjeev Arora, László Babai, Jacques Stern, and Z. Sweedyk. The hardness of approximate optima in lattices, codes, and systems of linear equations. *J. Comput. Syst. Sci.*, 54(2):317–331, 1997. Preliminary version in FOCS 1993.
- [2] Per Austrin, Subhash Khot, and Muli Safra. Inapproximability of vertex cover and independent set in bounded degree graphs. *Theory of Computing*, 7(1):27–43, 2011. Preliminary version in CCC 2009.
- [3] Tanmoy Chakraborty, Julia Chuzhoy, and Sanjeev Khanna. Network design for vertex connectivity. In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing (STOC)*, pages 167–176, 2008.
- [4] Siu On Chan. Approximation resistance from pairwise independent subgroups. *Electronic Colloquium on Computational Complexity (ECCC)*, 19:110, 2012.
- [5] Chandra Chekuri and Nitish Korula. Single-sink network design with vertex connectivity requirements. In *FSTTCS*, pages 131–142, 2008.
- [6] Chandra Chekuri and Nitish Korula. A graph reduction step preserving element-connectivity and applications. In *ICALP*, pages 254–265, 2009.
- [7] Joseph Cheriyan, Bundit Laekhanukit, Guylain Naves, and Adrian Vetta. Approximating rooted steiner networks. In *SODA*, pages 1499–1511, 2012.
- [8] Julia Chuzhoy and Sanjeev Khanna. Polynomial flow-cut gaps and hardness of directed cut problems. In *STOC*, pages 179–188, 2007.
- [9] Julia Chuzhoy and Sanjeev Khanna. Algorithms for single-source vertex connectivity. In *Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 105–114, 2008.
- [10] Julia Chuzhoy and Sanjeev Khanna. An $O(k^3 \log n)$ -approximation algorithm for vertex-connectivity survivable network design. In *Proceedings of the 50th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 437–441, 2009.
- [11] Julia Chuzhoy, Yury Makarychev, Aravindan Vijayaraghavan, and Yuan Zhou. Approximation algorithms and hardness of the k -route cut problem. In *SODA*, pages 780–799, 2012.
- [12] Reinhard Diestel. *Graph Theory (Graduate Texts in Mathematics)*. Springer, August 2005.
- [13] Michael Dinitz, Guy Kortsarz, and Ran Raz. Label cover instances with large girth and the hardness of approximating basic k -spanner. *ICALP 2012, To appear*, 2012. Also, available at <http://arxiv.org/abs/1203.0224>.
- [14] Irit Dinur and Prahladh Harsha. Composition of low-error 2-query pcps using decodable pcps. In *FOCS*, pages 472–481, 2009.
- [15] Oded Goldreich and Madhu Sudan. Locally testable codes and pcps of almost-linear length. *J. ACM*, 53(4):558–655, 2006. Preliminary version in FOCS 2002.

- [16] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bull. Amer. Math. Soc. (N.S.)*, 43(4):439–561 (electronic), 2006.
- [17] Subhash Khot and Muli Safra. A two prover one round game with strong soundness. In *FOCS*, pages 648–657, 2011.
- [18] Guy Kortsarz. On the hardness of approximating spanners. *Algorithmica*, 30(3):432–450, 2001.
- [19] Bundit Laekhanukit. An improved approximation algorithm for minimum-cost subset k -connectivity - (extended abstract). In *ICALP (I)*, pages 13–24, 2011.
- [20] Bundit Laekhanukit. An improved approximation algorithm for the minimum-cost subset k -connected subgraph problem. *CoRR*, abs/1104.3923, 2011. Preliminary version in [19].
- [21] Yuval Lando and Zeev Nutov. Inapproximability of survivable networks. *Theor. Comput. Sci.*, 410(21-23):2122–2125, 2009. Preliminary in APPROX 2008.
- [22] Dana Moshkovitz and Ran Raz. Two-query pcp with subconstant error. *J. ACM*, 57(5), 2010. Preliminary version in FOCS 2008.
- [23] Zeev Nutov. An almost $O(\log k)$ -approximation for k -connected subgraphs. In *SODA*, pages 912–921, 2009.
- [24] Zeev Nutov. Approximating minimum cost connectivity problems via uncrossable bifamilies and spider-cover decompositions. In *FOCS*, pages 417–426, 2009.
- [25] Zeev Nutov. A note on rooted survivable networks. *Inf. Process. Lett.*, 109(19):1114–1119, 2009. Preliminary version in [23].

A From Maximum To Minimum-Cost Label Cover

In this section, we show how to obtain the hardness of the minimum-cost label cover problem from the hardness of the maximum label cover problem.

The following is a standard lemma that transforms the hardness of the maximization version of the label cover problem to the hardness of the minimum-cost label cover problem. The theorem has been proved for the case that an instance has perfect completeness; see [1, 18], and also see [8, 3]. For our purpose, we state the theorem for the case of imperfect completeness.

Lemma 10. *Suppose there are constants $0 < \gamma, \epsilon < 1$ and $q_1, q_2 > 0$ such that, given an instance $(G = (U, W, E), \{\pi_e : e \in E\}, L_1, L_2)$ of the maximum label cover problem with $\Delta_{\text{left}} = q_1 \Delta_{\text{left}}^{\text{avg}}$ and $\Delta_{\text{right}} = q_2 \Delta_{\text{right}}^{\text{avg}}$, it is hard to distinguish between the following two cases:*

- **COMPLETENESS:** *There is a labeling covering at least $1 - \epsilon$ fraction of the arcs.*
- **SOUNDNESS:** *There is no labeling covering at least γ fraction of the arcs.*

Then it is hard to approximate the minimum-cost label cover problem to within a factor of $o(1/\sqrt{\gamma})$.

Proof. We construct an instance of the minimum-cost label cover problem from an instance $(G = (U, W, E), \{\pi_e : e \in E\}, L_1, L_2)$ of the maximum label cover problem with a parameter ϵ such that $\epsilon \cdot |E| \leq \min\{|U|, |W|\}$ as follows. First, we take an instance $(G = (U, W, E), \{\pi_e : e \in E\}, L_1, L_2)$ as a base construction. Then we set costs c_1 and c_2 of the left and right labels so that $c_1|U| = c_2|W|$, and let $C = c_1|U| + c_2|W|$. To show the hardness of the minimum-cost label cover problem, it suffices to show that there is a gap of at least $\sqrt{\gamma}/(16\sqrt{2q_1q_2}) = \Omega(\sqrt{\gamma})$ between the two cases of maximum label cover instances.

Completeness: Suppose there is a labeling (f_1, f_2) of the maximum label cover instance that covers $1 - \epsilon$ fraction of the arcs. Then, clearly, there is a labeling (\hat{f}_1, \hat{f}_2) of the minimum-cost label cover that covers the same number of arcs. For each arc (u, w) not covered, we add to $\hat{f}_1(u)$ and $\hat{f}_2(w)$ labels $a \in L_1$ and $b \in L_2$ such that $\pi_{u,w}(a) = b$. By the construction, the labeling (\hat{f}_1, \hat{f}_2) covers all the arcs. Since $\epsilon|E| \leq \min\{|U|, |W|\}$, the cost of the labeling (\hat{f}_1, \hat{f}_2) is at most $2(c_1|U| + c_2|W|) = 2C$.

Soundness: Suppose there is no labeling of the maximum label cover instance that covers at least $\gamma|E|$ arcs. We will show that if there is a labeling (\hat{f}_1, \hat{f}_2) of the minimum cost label cover instance with cost $\alpha C \leq (\sqrt{\gamma}/(8\sqrt{2q_1q_2})) \cdot C$, then there is a labeling (f_1, f_2) of the maximum label cover instance that covers at least $\gamma|E|$ arcs.

First, we construct (f_1, f_2) from (\hat{f}_1, \hat{f}_2) by uniformly at random picking a label $a \in \hat{f}_1(u)$ and assigning $f_1(u) = a$, for each $u \in U$, and uniformly at random picking a label $b \in \hat{f}_2(w)$ and assigning $f_2(w) = b$, for each $w \in W$. We claim that (f_1, f_2) covers at least $\gamma|E|$ arcs. To see this, consider the number of labels assigned to (\hat{f}_1, \hat{f}_2) . Let $U' \subseteq U$ and $W' \subseteq W$ be sets of vertices with at most $8\alpha q_1$ and $8\alpha q_2$ labels, respectively, and let $E' \subseteq E$ be the set of arcs with both endpoints in $U' \cup W'$. Then we have

$$|U - U'| \leq \frac{\alpha C}{8\alpha \cdot q_1 c_1} = \frac{1}{4q_1}|U| \text{ and } |W - W'| \leq \frac{\alpha C}{8\alpha \cdot q_2 c_2} = \frac{1}{4q_2}|W|$$

Thus, by union bound, the number of arcs of E' is at least

$$|E| - \frac{1}{4q_1}\Delta_{\text{left}}|U| - \frac{1}{4q_2}\Delta_{\text{right}}|W| \geq |E| - \frac{|E|}{2}$$

The second inequality follows from the facts that $\Delta_{\text{left}}|U| = q_1\delta_{\text{left}}|U| \leq q_1|E|$ and $\Delta_{\text{right}}|W| = q_2\Delta_{\text{right}}|W| \leq q_2|E|$. The probability that (f_1, f_2) covers any arc $e \in E'$ is at least $1/(64\alpha^2)$. Thus, the expected number of arcs of E' covered by (f_1, f_2) is

$$\sum_{e \in E'} (1 \cdot \Pr[(f_1, f_2) \text{ covers } e]) \geq \frac{|E|}{2} \cdot \frac{1}{64\alpha^2 \cdot q_1 q_2} \geq \frac{\gamma|E|}{128q_1 q_2 / (8\sqrt{2q_1 q_2})^2} = \gamma|E|.$$

We can derandomize this process by the method of conditional expectation. Therefore, there is a labeling (f_1, f_2) of the maximum label cover instance that covers at least $\gamma|E|$ arcs, a contradiction. \square

A similar lemma can be proven for the case of the label cover problem with perfect completeness. We will skip the proof for this case since it is almost identical to the previous one.

Lemma 11. *Suppose there are constants $0 < \gamma < 1$ and $q_1, q_2 > 0$ such that, given an instance $(G = (U, W, E), \{\pi_e : e \in E\}, L_1, L_2)$ of the maximum label cover problem with $\Delta_{\text{left}} = q_1\delta_{\text{left}}$ and $\Delta_{\text{right}} = q_2\Delta_{\text{right}}$, it is hard to distinguish between the following two cases:*

- **COMPLETENESS:** *There is a labeling covering all the arcs.*
- **SOUNDNESS:** *There is no labeling covering at least γ fraction of the arcs.*

Then it is hard to approximate the minimum-cost label cover problem to within a factor of $o(1/\sqrt{\gamma})$.

B The Right Degree Reduction

In this section, we discuss the right degree reduction introduced by Moshkovitz and Raz in [22]. The right degree reduction is an operation that transforms any instance of the maximum label cover problem to an instance with regular right degree d while preserving the completeness and preserving the soundness up to additive $O(1/\sqrt{d})$.

The right degree reduction is described as follows. Take an instance $(G = (U, W, E), \{\pi_e : e \in E\}, L_1, L_2)$ of the maximum label cover problem with completeness $1 - \epsilon$ and soundness γ . For each right vertex $w \in W$, we construct an expander graph H_w (H_w is an undirected graph) on $\deg(w)$ vertices with regular degree d and a second eigenvalue $O(\sqrt{d})$. Then we replace each vertex $w \in W$ by vertices of H_w . To be precise, we make $\deg(w)$ copies of w , namely $w(1), w(2), \dots, w(\deg(w))$, and associate each vertex $w(j)$ to a vertex of H_w by a one-to-one mapping. We order neighbors of w in G arbitrary, and let $u_1, u_2, \dots, u_{\deg(w)}$ be the neighbors of w . For each edge $\{w(i), w(j)\}$ of H_w , we add an arc $(u_i, w(j))$ and place a constraint $\pi_{u_i, w_j} = \pi_{u_i, w_j}$ on the arc (u_i, w_j) .

By the construction, there are d copies of arcs $e \in E$ in the output instance, and they have the same constraint. Thus, for any labeling (f_1, f_2) that covers $(1 - \epsilon)$ fraction of arcs of the input instance, there is a labeling (f_1, f'_2) that covers $(1 - \epsilon)$ fraction of arcs of the output instance, where f'_2 can be constructed by assigning $f'_2(w(j)) = f_2(w)$ for all copies $w(j)$ of a vertex $w \in W$. It was shown in [22] using the expander mixing lemma that the right-degree reduction gives an output instance with soundness $\gamma + O(1/\sqrt{d})$; see [22] and [14] for more detail.

This operation requires the projection property of a label cover instance and thus does not apply to the more general instance in which the constraints π_e are relations rather than projections. Also, the additive loss $O(1/\sqrt{d})$ in the soundness is the best possible because the smallest possible second eigenvalue of the d -regular expander graph is $2\sqrt{d} - 1$ due to the work of Alon and Boppana; see Theorem 5.3 in [16].

C Rooted k -Connectivity on Directed Graphs

In this section, we present hardness of $\Omega(k^{1/2})$ and $\Omega(\mathcal{D}^{1/4})$ for the rooted k -connectivity problem on directed graphs.

C.1 Hardness in Terms of k

First, we give a hardness of $\Omega(k^{1/2})$ for the rooted k -connectivity problem on directed graphs. Our result is based on the construction in [7]. Here we will give a construction but will omit the proof. Let $(G = (U, W, E), \{\pi_e : e \in E\}, L_1, L_2, c_1, c_2)$ be an instance of the minimum-cost label cover problem. We construct a directed graph $\hat{G} = (\hat{V}, \hat{E})$ of the rooted k -connectivity problem as follows.

Base Construction: For each vertex $u_i \in U$, we add to \hat{G} a vertex u_i and a set of vertices A_i , which is a copy of the set of labels L_1 ; we join u_i to each vertex $a \in A_i$ by an arc (u_i, a) . For each vertex $w_j \in W$, we add to \hat{G} a vertex w_j and a set of vertices B_j , which is a copy of the set of labels L_2 ; we join w_j to each vertex $b \in B_j$ by an arc (w_j, b) . We may think that u_i (resp., w_j) is the same vertex in both G and \hat{G} . Also, since A_i (resp., B_j) is a copy of L_1 (resp., L_2), we may say that a vertex $a \in A_i$ (resp., $b \in B_j$) is a label in L_1 (resp., L_2). We set cost c_1 on an arc (u_i, a) , for each $a \in A_i$, and we set cost c_2 on an arc (w_j, b) , for each $b \in B_j$. For each arc (u_i, w_j) of G , we add to \hat{G} a zero-cost arc (a, b) joining a vertex $a \in A_i$ to a vertex $b \in B_j$ if $\pi_{u_i, w_j}(a) = b$. This finishes the base construction.

The final construction: Now, we add a root vertex r to \hat{G} and join r to each vertex u_i by a zero-cost arc (r, u_i) . For each arc (u_i, w_j) , we add a terminal $t_{i,j}$ and join w_j to $t_{i,j}$ by a zero-cost arc $(w_j, t_{i,j})$. Thus, we have the root vertex r and a set of terminals $T_{i,j} = \{t_{i,j} : (u_i, w_j) \in E\}$. Next, we add a zero-cost arcs $(u_{i'}, t_{i,j})$, called a *padding arc*, if $i' \neq i$ and $(u_{i'}, w_j) \in E$. Thus, each terminal has indegree at most $\Delta(G)$. For each terminal $t_{i,j}$ with indegree $d_{i,j} < \Delta(G)$, we add $\Delta(G) - d_{i,j}$ copies of a zero-cost arc $(r, t_{i,j})$. Finally, we set the connectivity requirement $k = \Delta(G)$. (Note that $\Delta(G) \ll |T|$.)

The above construction gives the following theorem whose correctness is proved in [7].

Theorem 12 ([7]). *There is a polynomial-time approximation preserving reduction such that, given an instance of the minimum-cost label cover problem consisting of a graph G , outputs an instance of the rooted k -connectivity problem on directed graphs with $k = \Delta(G)$.*

Applying Theorem 3, it then immediately follows that the hardness of the rooted k -connectivity problem on directed graphs is $\Omega(k^{1/2-\epsilon})$, for any $\epsilon > 0$ (since $k = \Delta(G)$). Thus, we have the next theorem.

Theorem 13. *For $k < |T|$, unless $\text{NP} = \text{ZPP}$, it is hard to approximate the rooted k -connectivity problem on directed graphs to within a factor of $o(k^{1/2})$.*

C.2 Hardness in Terms of \mathcal{D}

Now, we show the hardness of the rooted k -connectivity problem on directed graphs in terms of the other parameter. Specifically, we show a hardness of $\Omega(\mathcal{D}^{1/4})$ for the rooted k -connectivity problem on directed graphs, where $\mathcal{D} = |T|$ (since the demand pairs are between the root vertex r and terminals in T). We start from the previous construction and then merge some terminals. The key idea is to merge terminals that do not share paths from the root vertex. To be precise, we say that two terminals $t_{i,j}$ and $t_{i',j'}$ are *dependent* if there are an $r, t_{i,j}$ -path P and an $r, t_{i',j'}$ -path P' that have a common vertex $v \neq r$; otherwise, we say that $t_{i,j}$ and $t_{i',j'}$ are *independent*. Observe that two terminals $t_{i,j}$ and $t_{i',j'}$ are dependent if and only if

arcs (u_i, w_j) and $(u_{i'}, w_{j'})$ of G (of the label cover instance) are incident or there is an arc joining them. Specifically, (u_i, w_j) and $(u_{i'}, w_{j'})$ are independent if and only if they form an induced matching in G . This proves in the lemma below. For notational convenience, we use ij to mean an arc (u_i, w_j) of G .

Lemma 14. *Any two terminals $t_{i,j}$ and $t_{i',j'}$ are independent if and only if ij and $i'j'$ forms an induced matching in G .*

Proof. First, we prove the “only if” part. Suppose $t_{i,j}$ and $t_{i',j'}$ are independent, but $\{ij, i'j'\}$ is not an induced matching in G . Then we have two cases: (1) $i = i'$ or $j = j'$ (2) G has an arc ij' or $i'j$. For the former case, every $r, t_{i,j}$ -path and $r, t_{i',j'}$ -path in G has to use either u_i or w_j and thus share a vertex. This implies that $t_{i,j}$ and $t_{i',j'}$ are dependent, a contradiction. For the latter case, assume wlog that G has an arc ij' . Then we must have added to G a padding arc $u_i, t_{i',j'}$ by the construction. Thus, there is an $r, t_{i',j'}$ -path that shares a vertex u_i with an $r, t_{i,j}$ -path, a contradiction. This proves the “only if” part.

Next, we prove the “if” part. Suppose ij and $i'j'$ form an induced matching in G . Then $i \neq i'$ and $j \neq j'$. Also, G has neither an arc ij' nor an arc $i'j$. It then follows immediately by the construction that \hat{G} has no $r, t_{i,j}$ -path and $r, t_{i',j'}$ -path that share a common vertex. This completes the proof. \square

Lemma 14 allows us to apply a strong edge coloring algorithm to the arcs of G , which are constraints of the label cover instance. It is known that every graph G can be strongly colored using at most $2\Delta(G)^2$ colors. Since each color class forms an induced matching in G , no two of them are dependent. Thus, we can merge all the terminals corresponding to arcs of the same color class into one terminal without any conflict. To be precise, for each color C , define $T_C = \{t_{i,j} : ij \text{ has color } C\}$. Then we unify T_C as a single terminal and set a connectivity requirement $k|T_C|$ for this terminal. The new graph is denoted by \hat{G}^{new} . Observe that any $k|T_C|$ openly disjoint r, T_C -paths in \hat{G}^{new} corresponds to k openly disjoint $r, t_{i,j}$ -paths for every $t_{i,j} \in T_C$ in the original graph \hat{G} . Thus, there is a one-to-one mapping between the solution in the new instance and that of the old instance, and both have the same hardness. To make a connectivity uniform, set $k^{new} = k \cdot \max_C |T_C|$ and add $k^{new} - k|T_C|$ copies of a zero-cost arc (r, T_C) for each terminal T_C . By the construction, we have at most $2\Delta(G)^2$ terminals in the new instance. Therefore, applying Theorem 3, we have a hardness of $\Omega(\mathcal{D}^{1/4-\epsilon}) = \Omega(|T|^{1/4-\epsilon})$ for the rooted k -connectivity problem on directed graphs, for any constant $\epsilon > 0$.

Theorem 15. *For $k \geq |T|$, unless $\text{NP} = \text{ZPP}$, it is hard to approximate the rooted k -connectivity problem on directed graphs to within a factor of $o(\mathcal{D}^{1/4-\epsilon}) = o(|T|^{1/4-\epsilon})$ for any constant $\epsilon > 0$.*

D The Rooted k -Connectivity Problem on Undirected Graphs.

In this section, we present hardness constructions for the rooted k -connectivity on undirected graphs.

D.1 Hardness in Terms of k

Similar to the case of the directed graphs, the following theorem has been proved in [7].

Theorem 16 ([7]). *There is a polynomial-time approximation preserving reduction such that, given an instance of the minimum-cost label cover problem consisting of a graph G with a set of labels L_1 and L_2 , outputs an instance of the rooted k -connectivity problem on directed graphs with $k = O(\Delta(G)^3 \cdot \max\{|L_1|, |L_2|\} + \Delta(G)^5)$.*

We will present the hardness construction described in [7] but will skip the proof for completeness and soundness. (For more detail, see [7].) Take an instance $(G = (U, W, E), \{\pi_e : e \in E\}, L_1, L_2, c_1, c_2)$ of the minimum-cost label cover problem. We construct a graph $\hat{G} = (\hat{V}, \hat{E})$ of the rooted k -connectivity problem on undirected graphs as follows.

Base Construction: The base construction is the same as that in Section C.1 except that we ignore the direction of edges.

Add Root Vertex and Terminals: We add to \hat{G} a root vertex r . For each arcs $(u_i, w_j) \in G$, we add to \hat{G} a clique $X_{i,j}$ and a terminal $t_{i,j}$; the size of $X_{i,j}$ will be specified later. We join each clique $X_{i,j}$ to a vertex $u_i \in U$ by adding a zero-cost edge $\{x, u_i\}$ for each pair of vertices $x \in X_{i,j}$ and $u_i \in U$. We join each terminal $t_{i,j}$ to a vertex $w_j \in W$ by a zero-cost edge $\{w_j, t_{i,j}\}$. Then we join the root vertex r to each clique $X_{i,j}$ by a zero-cost edge $\{r, x\}$ for each $x \in X_{i,j}$.

Final Construction: Now, we add some zero-cost edges, called *padding edges*, which intuitively force $r, t_{i,j}$ -paths to be in a *canonical* form. We say that an $r, t_{i,j}$ -paths is a *canonical path* if it is of the form $r, X_{i,j}, u_i, A_i, B_j, w_j, t_{i,j}$. The padding for each terminal $t_{i,j}$ is as follows. For notational convenience, we use ij to mean an arc $\{u_i, w_j\}$ of the label cover instance G , and we use $\text{dist}(ij, i'j')$ to mean the distance between ij and $i'j'$ in the *line graph* H of the underlying undirected of G , i.e., the vertex set of H is the edge set of G , and there is an edge (e, e') in H if edges e and e' share an endpoint in G . We define the set $Z_{i,j}$ and $Y_{i,j}$ as below.

$$\begin{aligned} Z_{i,j}^1 &= \left(\bigcup_{i' \neq i: i'j \in E} A_{i'} \right) \cup \left(\bigcup_{j' \neq j: ij' \in E} B_{j'} \right) \\ Z_{i,j}^2 &= \{t_{i',j'} : 1 \leq \text{dist}(ij, i'j') \leq 2\} \\ Z_{i,j} &= Z_{i,j}^1 \cup Z_{i,j}^2 \\ Y_{i,j} &= \bigcup_{1 \leq \text{dist}(ij, i'j') \leq 2} X_{i',j'} \end{aligned}$$

We also create a set of vertices $Q_{i,j}$, which is the set of auxiliary vertices created to make the connectivity requirement uniform. The vertices of $Q_{i,j}$ are not in the base construction, and its size will be specified later. We join $Y_{i,j}$ to $t_{i,j}$ by edges $\{y, t_{i,j}\}$ for all $y \in Y_{i,j}$. We join $Z_{i,j}$ to $X_{i,j}$ and $t_{i,j}$ by edges $\{x, z\}, \{z, t_{i,j}\}$ for all pairs of vertices $x \in X_{i,j}$ and $z \in Z_{i,j}$. We join $r, Q_{i,j}$ and $t_{i,j}$ by edges $\{r, q\}, \{q, t_{i,j}\}$ for all $q \in Q_{i,j}$. All of these edges have zero costs.

Lastly, we have to set the connectivity requirement and specify the size of $X_{i,j}$. We remark that we want all the $r, t_{i,j}$ -paths except a canonical path to use all the vertices in $Z_{i,j} \cup Y_{i,j} \cup Q_{i,j}$. Thus, we need to set the size of $X_{i,j}$ to be $|Z_{i,j}| + 1$. By the construction, we have to set $|Q_{i,j}| = k - |Z_{i,j}| + |Y_{i,j}| - 1$.

$$|Z_{i,j}| = |Z_{i,j}^1| + |Z_{i,j}^2| \leq 2\Delta(G) \cdot \max\{|L_1|, |L_2|\} + 2\Delta(G)^2.$$

Thus, we need $|X_{i,j}| = 2\Delta(G) \cdot \max\{|L_1|, |L_2|\} + 2\Delta(G)^2 + 1$, implying that $|Y_{i,j}| = O(\Delta(G)^3 \cdot \max\{|L_1|, |L_2|\} + \Delta(G)^4)$. We set $k = \max_{i,j} (|Z_{i,j}| + |Y_{i,j}|) + 1$ and set the size of $Q_{i,j}$ to be $|Q_{i,j}| = k - |Z_{i,j}| + |Y_{i,j}| - 1$. Therefore, we have an instance of the rooted k -connectivity problem on undirected graphs with $k = O(\Delta(G)^3 \cdot \max\{|L_1|, |L_2|\} + \Delta(G)^4)$ (given an instance in Theorem 3, we have $k = O(q^5)$), and the reduction preserves both the completeness and soundness. Here we skip the completeness and soundness proofs. For more detail, please see [7].

D.2 Hardness in terms of \mathcal{D}

To obtain the hardness in terms of \mathcal{D} for the rooted k -connectivity problem on undirected graphs, we may apply the same technique as that used in the directed case. However, we can simplify the proof by applying the following theorem due to Lando and Nutov [21].

Theorem 17 ([21]). *There is a polynomial-time approximation preserving reduction that, given an instance of the “directed” rooted k -connectivity problem consisting of a directed graph G on n vertices, a root vertex r , a set of terminals T and a connectivity requirement k , outputs an instance of the “undirected” rooted k -connectivity problem consisting of an undirected graph G' on $n' = 2n$ vertices, a root vertex r , and a set of terminals T' , where $|T'| = |T|$ and $k' = k + n$.*

Since there is an $\Omega(|T|^{1/4-\epsilon})$ -hardness for the rooted k -connectivity problem on directed graphs, the same hardness applies for the undirected case as well.

Theorem 18. *For $k \geq |T|$, it is NP-hard to approximate the undirected rooted k -connectivity problem to within a factor of $o(|T|^{1/4-\epsilon}) = o(\mathcal{D}^{1/4-\epsilon})$ for any constant $\epsilon > 0$.*

E Vertex-Connectivity Survivable Network Design

In this section, we present hardness constructions of the vertex-connectivity survivable network design problem on undirected graphs.

E.1 Hardness in Terms of k

The hardness of the vertex-connectivity survivable network design problem can be derived from its special case, the rooted k -connectivity problem. However, by applying the reduction directly from the minimum-cost label cover problem, we have a better bound.

The following theorem is proved by Chakrabarty, Chuzhoy and Khanna in [3].

Theorem 19 ([3]). *There is a polynomial-time approximation preserving reduction such that, given an instance of the minimum-cost label cover problem consisting of a graph G with a set of labels L_1 and L_2 , outputs an instance of the rooted k -connectivity problem on directed graphs with $k = O(\Delta(G) \cdot \max\{|L_1|, |L_2|\} + \Delta(G)^2)$.*

We will present the hardness construction described in [3] but will skip the proof for completeness and soundness. (For more detail, see [3].) Take an instance $(G = (U, W, E), \{\pi_e : e \in E\}, L_1, L_2, c_1, c_2)$ of the minimum-cost label cover problem. We construct a graph $\widehat{G} = (\widehat{V}, \widehat{E})$ of the vertex-connectivity survivable network design problem on undirected graphs as follows.

Base Construction: The base construction is the same as that in Section C.1 except that we ignore the direction of edges.

Add Source-Sink Pairs: We will add to the undirected graph \widehat{G} *source-sink* pairs, i.e., we add a pair of vertices whose connectivity requirement is positive. For each arcs $(u_i, w_j) \in G$, we add to \widehat{G} a source $s_{i,j}$ and a sink $t_{i,j}$. We join each source $s_{i,j}$ to a vertex $u_i \in U$ by adding a zero-cost edge $\{s_{i,j}, u_i\}$, and we join each terminal $t_{i,j}$ to a vertex $w_j \in W$ by a zero-cost edge $\{w_j, t_{i,j}\}$.

Final Construction: Now, we add some zero-cost edges, called *padding edges*, which intuitively force $s_{i,j}, t_{i,j}$ -paths to be in a *canonical* form. We say that an $r, t_{i,j}$ -paths is a *canonical path* if it is of the form

$s_{i,j}, u_i, A_i, B_j, w_j, t_{i,j}$. The padding for a source-sink pair $s_{i,j}, t_{i,j}$ is as follows. For notational convenience, we use ij to mean an arc $\{u_i, w_j\}$ of the label cover instance G , and we use $\text{dist}(ij, i'j')$ to mean the distance between ij and $i'j'$ in the line graph H of the underlying undirected of G , i.e., the vertex set of H is the edge set of G , and there is an edge (e, e') in H if edges e and e' share an endpoint in G . We define the set $Z_{i,j}$ and $Y_{i,j}$ as below.

$$Z_{i,j} = \left(\bigcup_{i' \neq i: i'j \in E} A_{i'} \right) \cup \left(\bigcup_{j' \neq j: ij' \in E} B_{j'} \right)$$

$$Y_{i,j} = \bigcup_{1 \leq \text{dist}(ij, i'j') \leq 2} \{s_{i',j'}, t_{i',j'}\}$$

We join $s_{i,j}$ and $t_{i,j}$ to $Y_{i,j} \cup Z_{i,j}$ by adding zero-cost edges $\{s_{i,j}, x\}$ and $\{x, t_{i,j}\}$ for all $x \in Y_{i,j} \cup Z_{i,j}$. For the connectivity requirement, we set $\text{req}(s_{i,j}, t_{i,j}) = |Y_{i,j} \cup Z_{i,j}| + 1$ for all source-sink pairs $s_{i,j}, t_{i,j}$. We may make the requirements uniform by setting $k = \max_{i,j} |Y_{i,j} \cup Z_{i,j}| + 1$ and adding a set of auxiliary vertices $Q_{i,j}$ with $|Q_{i,j}| = k - |Y_{i,j} \cup Z_{i,j}| - 1$ for each source-sink pair s_i, t_j .

By the construction, we have

$$|Z_{i,j}| + |Y_{i,j}| \leq 2\Delta(G) \cdot \max\{|L_1|, |L_2|\} + 4\Delta(G)^2$$

Therefore, we have an instance of the rooted k -connectivity problem on undirected graphs with $k = O(\Delta(G) \cdot \max\{|L_1|, |L_2|\} + \Delta(G)^2)$ (given an instance in Theorem 3, we have $k = O(q^3)$), and the reduction preserves both the completeness and soundness. Here we skip the completeness and soundness proofs. For more detail, please see [3].

E.2 Hardness in Terms of \mathcal{D}

The hardness in terms of \mathcal{D} for the vertex-connectivity survivable network design problem on undirected graphs follows immediately from that of its special case, the rooted k -connectivity problem on undirected graphs. Thus, we have

Theorem 20. *For $k \geq \mathcal{D}$, it is NP-hard to approximate the vertex-connectivity survivable network design problem on undirected graphs to within a factor of $o(\mathcal{D}^{1/4-\epsilon})$ for any constant $\epsilon > 0$.*

F Vertex-Connectivity k -Route Cut

In this section, we discuss the vertex-connectivity k -route cut problem.

F.1 Hardness in Terms of k

The following theorem is proved in [11] by Chuzhoy et al.

Theorem 21 ([11]). *There is a polynomial-time approximation preserving reduction such that, given an instance of the minimum-cost label cover problem consisting of a graph G with a set of labels L_1 and L_2 , outputs an instance of the vertex-connectivity k -route cut problem on undirected graphs with $k = O(\Delta(G) \cdot \max\{|L_1|, |L_2|\})$.*

We will give the hardness construction of this problem based on the construction in [11]³. Take an instance $(G = (U, W, E), \{\pi_e : e \in E\}, L_1, L_2, c_1, c_2)$ of the minimum-cost label cover problem. We construct a graph $\widehat{G} = (\widehat{V}, \widehat{E})$ of the vertex-connectivity k -route cut problem as follows.

Base Construction: First, for each left vertex $u_i \in U$, we create a set of edges $E(u_i) = \{\{a, a'\} : a \in L_1\}$. Similarly, for each right vertex $w_j \in W$, we create a set of edges $E(w_j) = \{\{b, b'\} : b \in L_2\}$. All the edges in $E(u_i)$'s have costs c_1 , and all the edges in $E(w_j)$'s have costs c_2 .

Next, for each right vertex $w_j \in W$, we arrange edges in $E(w_j)$ in an arbitrary order, say $E(w_j) = \{\{b_1, b'_1\}, \{b_2, b'_2\}, \dots, \{b_{|L_2|}, b'_{|L_2|}\}\}$. We then form a path P_j by joining edges in $E(w_j)$ by edges with costs infinity. To be precise, we have an edge $\{b_\ell, b_{\ell+1}\}$ with cost infinity in \widehat{G} , for each $\ell = 1, 2, \dots, |L_2| - 1$. Because of the projection property, which we will discuss later, we only create such paths for right vertices.

For each arc $(u_i, w_j) \in E$, we construct a path $Q_{i,j}$ as follows. For any label $b_\ell \in L_2$, let $\pi^{-1}(b_\ell) = (a_{q_1}, a_{q_2}, \dots, a_{q_{\alpha(b_\ell)}})$ be a sequence of labels in L_1 that projects to b_ℓ arranged in an arbitrary order. Since each label a_{q_t} maps to an edge $\{a_{q_t}, a'_{q_t}\}$ in $E(w_j)$, we may abuse $\pi^{-1}(b)$ to mean edges in $E(u_i)$. We define an order $\psi(i, j)$ by arranging the edges of $E(u_i)$ to be of the form

$$\pi^{-1}(b_1), \pi^{-1}(b_2), \dots, \pi^{-1}(b_{|L_2|})$$

Then we form a path $Q_{i,j}$ by joining edges in $E(u_i)$ according to the order $\psi(i, j)$ by paths of length 2. That is,

$$Q_{i,j} = (a_1, a'_1, x_1, a_2, a'_2, x_2, \dots, a_{|L_1|}, a'_{|L_1|})$$

where indices of a_q 's are obtained from $\psi(i, j)$. (This is crucial as edges in $E(u_i)$ may have different orders in two different paths $Q_{i,j}$ and $Q_{i,j'}$.) We denote by $X_{i,j}$ a set of vertices $x'_{q's}$ in $Q_{i,j}$ separating edges of $E(u_i)$. Next, for each edge $\{b_\ell, b'_\ell\}$ and $\{b_{\ell+1}, b'_{\ell+1}\}$ in $E(w_j)$, we join b'_ℓ and $b_{\ell+1}$ to a vertex x_{q_ℓ} , where x_{q_ℓ} is a vertex in $X_{i,j}$ connecting the path on $\pi^{-1}(b_\ell)$ to the path on $\pi^{-1}(b_{\ell+1})$. These edges have costs infinity. This completes the base construction.

Add Source-Sink Pairs: For each vertex $u_i \in U$, we add a source vertex s_i , and for each vertex $w_j \in W$, we add a sink vertex t_j . For each arc (u_i, w_j) of G , we add a demand (source-sink) pair (s_i, t_j) to the set of demand pairs \mathbb{D} ; then we join s_i to the first vertices of the paths $Q_{i,j}$ and P_j , and we join t_j to the last vertices of the paths $Q_{i,j}$ and P_j . All of these edges have costs infinity.

Final Construction: Now, we add *padding edges*, which will guarantee that to make the vertex-connectivity of a source-sink pair s_i, t_j to be below k , ones have to remove edges corresponding to the feasible labeling of the label cover instance. For the ease of presentation, we will use ij to mean an arc (u_i, w_j) of G . Denote by $V(F)$ a set of vertices spanned by a set of edges (resp., a graph) F . We define a set of vertices $Z_{i,j}$ to be the set of neighbors of $V(Q_{i,j}) \cup V(P_j)$ in the current graph; that is,

$$Z_{i,j} = \bigcup_{i'j' \in E(G): i' \neq i} (V(Q_{i',j'}) \cup \{s_{i'}\}) \cup \bigcup_{ij' \in E(G): j' \neq j} (X_{i,j'} \cup V(P_{j'}) \cup \{t_{j'}\})$$

By the construction, we have

$$z = \max_{ij \in E} |Z_{i,j}| \leq O(\Delta(G)(|L_1| + |L_2|)) \leq O(\Delta(G) \cdot \max\{|L_1|, |L_2|\}).$$

³ Due to a very subtle error in the proof in [11], our construction is slightly different from the original construction.

The size $Z_{i,j}$ and $Z_{i',j'}$ for $ij \neq i'j'$ may be different. Thus, we add to \widehat{G} a set of auxiliary vertices $S_{i,j}$ with $|S_{i,j}| = z - |Z_{i,j}|$. Then we add edges $\{s_i, v\}, \{t_j, v\}$ with cost infinity joining s_i and t_j to each vertex $v \in Z_{i,j} \cup S_{i,j}$. Note that all the neighbors of $\{s_i, t_j\} \cup V(Q_{i,j}) \cup V(P_j)$ are in $Z_{i,j} \cup S_{i,j}$. Finally, we set $k = z + 1$, finishing the construction.

Completeness and Soundness: Both the completeness and soundness proofs follows from the next claim.

Claim 22. *Consider any source-sink pair $(s_i, t_j) \in \mathbb{D}$. There are at most $k + 1$ openly disjoint s_i, t_j -paths in \widehat{G} , and there are at most 2 openly disjoint s_i, t_j -paths in $\widehat{G} - (Z_{i,j} \cup S_{i,j})$. Moreover, any 2 openly disjoint s_i, t_j -paths in $\widehat{G} - (Z_{i,j} \cup S_{i,j})$ must have $Q_{i,j}$ and P_j as subpaths.*

Proof. First, observe that vertices in $Z_{i,j} \cup S_{i,j}$, $V(Q_{i,j})$ and $V(P_j)$ are pairwise disjoint by the construction. It can be seen that $Z_{i,j} \cup S_{i,j}$ gives $k - 1$ paths between s_i and t_j , and the path $Q_{i,j}$ and P_j give another two paths. Thus, we proved both the first and the second statements. For the third statement, it follows by the construction that all the neighbors of $\{s_i, t_j\} \cup V(Q_{i,j}) \cup V(P_j)$ are in $Z_{i,j} \cup S_{i,j}$, which means that the only way we can have 2 openly disjoint s_i, t_j -paths in $\widehat{G} - (Z_{i,j} \cup S_{i,j})$ is to follow the two paths $Q_{i,j}$ and P_j . Hence, the claim follows. \square

It can be seen that there is a one-to-one mapping between edges with finite cost and the labels of the minimum-cost label cover instance. Thus, it suffices to show that, for any edges $E' \subseteq E$ with finite cost, $G - E'$ has no k openly disjoint s_i, t_j paths for all $(s_i, t_j) \in \mathbb{D}$ if and only if the labeling (f_1, f_2) corresponding to E' is feasible for the minimum-cost label cover instance. To see this, consider the graph $\widehat{G}_{i,j} = \widehat{G} - (Z_{i,j} \cup S_{i,j})$ for any source-sink pair $(s_i, t_j) \in \mathbb{D}$. By Claim 22, there are at most 2 openly disjoint s_i, t_j -paths, and these paths have to use $Q_{i,j}$ and P_j . Now, consider the paths $Q_{i,j}$ and P_j . Observe that if we remove an edge $\{b, b'\}$ and one edge in π_{u_i, w_j}^{-1} , then the resulting graph has no s_i, t_j -path. Conversely, if we remove $\{b, b'\}$ but none of the edges in π_{u_i, w_j}^{-1} , then we can have a path that goes zig-zag between $Q_{i,j}$ and P_j via an edge $\{x_q, \hat{b}\}$, where $\{\hat{b}, \hat{b}'\}$ is an edge next to $\{b, b'\}$. Therefore, we conclude that there is no k openly disjoint s_i, t_j -paths in $\widehat{G}_{i,j}$ if and only if we remove edges corresponding to the labeling covering an edge $\{u_i, w_j\}$ of G . This completes the proof of Theorem 21. \square

F.2 Hardness in Terms of \mathcal{D}

For the hardness in terms of the number of demand pairs, it can be seen that for any pair of arcs (u_i, w_j) and $(u_{i'}, w_{j'})$ that form an induced matching (i.e., G has no arc joining (u_i, w_j) and $(u_{i'}, w_{j'})$), all the vertices and edges of subgraphs of \widehat{G} induced by $Z_{i,j} \cup S_{i,j} \cup V(Q_{i,j}) \cup V(P_j)$ and $Z_{i',j'} \cup S_{i',j'} \cup V(Q_{i',j'}) \cup V(P_{j'})$ are disjoint by the construction. Thus, two arcs (u_i, w_j) and $(u_{i'}, w_{j'})$ are *independent* if and only if (u_i, w_j) and $(u_{i'}, w_{j'})$ form an induced matching. By the same arguments as in the case of the rooted k -connectivity problem on directed graphs, we have the following theorem.

Theorem 23. *For $k \geq \mathcal{D}$, unless $\text{NP} = \text{ZPP}$, it is hard to approximate the vertex connectivity k -route cut problem to within a factor of $o(\mathcal{D}^{1/4-\epsilon})$ for any constant $\epsilon > 0$.*